Estuarine Management and Technologies Journal

Estuarine Management and Technologies 1: 57–68 (2024)
DOI: 10.3897/emt.1.130278

Research Article

# Predicting chlorophyll-a dynamics in the Ashtamudi estuary using ANN and ANFIS techniques

K. L. Priya[1], Soufiane Haddout[2], Joan Cecilia C. Casila[3]

1 *Department of Civil Engineering, TKM College of Engineering, Kollam, India*
2 *Department of Physics, Faculty of Science, Ibn Tofail University, Kenitra, Morocco*
3 *Land and Water Resources Engineering Division, Institute of Agricultural and Biosystems Engineering, College of Engineering and Agro-Industrial Technology, University of the Philippines, Los Baños, Laguna, Philippines*

*Corresponding author: K. L. Priya (priyaram@tkmce.ac.in, klpriyaram@gmail.com)*

OPEN ACCESS

Academic editor: Hermano Melo Queiroz
**Received:** 22 June 2024
**Accepted:** 2 October 2024
**Published:** 21 October 2024

**Citation:** Priya KL, Haddout S, Casila JCC (2024) Predicting chlorophyll-a dynamics in the Ashtamudi estuary using ANN and ANFIS techniques. Estuarine Management and Technologies 1: 57–68. https://doi.org/10.3897/emt.1.130278

**Copyright:** © K. L. Priya et al.
This is an open access article distributed under terms of the Creative Commons Attribution License (Attribution 4.0 International – CC BY 4.0).

## Abstract

Excessive nutrients and associated high chlorophyll-a concentration are of growing concern to the public health due to the occurrence of eutrophication in aquatic bodies. Continuous monitoring of chlorophyll-a has limitations, thereby necessitating the development of prediction models. Even though chlorophyll-a prediction models are numerous, they rarely fit into estuarine conditions, which is encountered with dynamic mixing of freshwater and sea water. Thus, the present study identifies the most appropriate driving factors that are relevant for estuarine conditions through chlorophyll-a prediction models developed using Artificial Neural Network and Adaptive Neuro Fuzzy Inference System (ANFIS) approaches in the Ashtamudi estuary, India. Salinity, which is an indirect measure of the estuarine mixing intensity, has been used to replicate the dynamic mixing in the estuary. The results indicate that the model incorporating salinity along with total nitrogen, total phosphorous and turbidity well-predicted chlorophyll-a concentration with a coefficient of determination ranging from 0.73 to 0.99. General Regression Neural Network (GRNN) and ANFIS models outperformed Back Propagation Neural Network (BPNN) and Recurrent Neural Network (RNN) models. Of the various ANFIS models, the ones that used Gaussian and Generalized Bell membership functions were most appropriate for the prediction of chlorophyll-a than the models with triangular and trapezoidal membership functions. The study identified that the models that considered salinity were less sensitive and can be used for modelling chlorophyll-a. The chlorophyll-a was observed to have a direct influence of salinity at salinity values greater than 5‰. The study highlighted that estuarine mixing further enhances the primary productivity through increased penetration of light leading to higher chlorophyll-a concentration. Further the study emplasized that the predictive models are important tools fitting well within estuarine environments due to its replicability.

**Key words:** ANFIS, ANN, Chlorophyll-a, estuary, eutrophication, prediction

## Introduction

Eutrophication has been recognized as a major problem in aquatic environments (Mozo et al. 2022). The excessive influx of nutrients such as nitrogen and phosphorous lead to algal blooms, thereby posing detrimental effects to the water quality, aquatic organism and successively affecting socio-economic and environmental

health. Chlorophyll-a is often considered as a proxy to eutrophication levels and it can indicate the trophic state of the aquatic body (Savoy and Harvey 2023). It can be used as a measure of the overall phytoplankton biomass, which represents the primary productivity of the ecosystem (Chang et al. 2021). Higher concentrations of chlorophyll-a in water can lead to bad odour and taste, DO depletion, enhanced turbidity and organic load. This can in turn influence fisheries, tourism, public health, thereby implying a major detrimental socio-economic impact. Therefore, monitoring chlorophyll-a concentration is highly essential for remediating eutrophication through developing appropriate management practices.

Determining chlorophyll-a on a daily basis is highly time-consuming and because of this, the monitoring agencies adopt a monthly sampling protocol. In order to overcome the difficulty in analysis, many researchers have developed prediction models in several aquatic environment. Developing models for chlorophyll-a is rather challenging as it can exhibit a non-linear behaviour with its driving factors (Savoy and Harvey 2023). Moreover, identifying the driving factors of chlorophyll-a is still more tedious as chlorophyll-a concentrations may depend on various site-specific factors. When compared to lentic ecosystems, turbulent mixing plays important role in lotic ecosystems like rivers and estuaries. The estuarine mixing is controlled by the river flow and tidal flow and a single easily accessible parameter that can represent the mixing conditions is the salinity values (Priya et al. 2015). Nevertheless, researchers have not accounted for salinity while modelling for chlorophyll-a in estuaries. This necessitates a detailed investigation of the influence of salinity on the prediction of chlorophyll-a by considering salinity as one of the input parameters.

Various studies have adopted linear regression, multiple linear regression, non-linear regression, and machine learning methods for modelling chlorophyll-a. Amorim et al. (2021) developed chlorophyll-a prediction models using Back propagation neural networks, random forest and support vector machine tools and observed that support vector machine algorithm marginally outperformed other models. LSTM neural network tool has been used by Cen et al. (2022) for predicting chlorophyll-a in the East China sea using 15 days chlorophyll-a data. However, they have not considered some of the drivers that represent estuarine mixing. Organic carbon content and DO were identified as major environmental factors in predicting chlorophyll-a using various neural network approaches in the coastal waters of Korea (Kim et al. 2022). Kim and Ahn (2022) identified total phosphorous, total nitrogen, pH, DO, EC and organic carbon as significant parameters for the prediction of chlorophyll-a using ANN techniques. Huang et al. (2023) has coupled Remote Sensing data with in-situ data and machine learning tools for predicting chlorophyll-a, while Song et al. (2024) identified turbidity and salinity as important factors affecting chlorophyll-a in the Yellow River estuary. From these studies, it is evident that the tools and variables for chlorophyll-a prediction are different for various types of environments and hence can be explained as site-specific. Further, salinity has been identified to influence chlorophyll-a (Gayathri et al. 2022) and hence incorporating salinity as an input variable can possibly enhance the predictability of models. Previous models have limitations of not considering estuarine mixing. Thus, the main aim of the study is to identify the significant driving parameters of chlorophyll-a and the appropriate ML model for the prediction of chlorophyll-a. Further, the influence of salinity, which is an indirect representation of mixing of freshwater and tidal water on chlorophyll-a

prediction has been evaluated. In order to test the predictive model performance, is to test a predictive model, the Ashtamudi estuary on the south-west coast of India has been selected as a framework due to its historical significance.

## 2. Methodology

### 2.1 Study Area

Ashtamudi estuary is one among the Ramsar sites and is renowned for its fish diversity. Located between 8°53'N to 9°2'N latitude and 76°31'E to 76°41'E longitude, the estuary is connected to the Arabian Sea through a perennial opening (Fig. 1). The estuary receives freshwater discharge from the Kallada River, which originates from the Kulathupuzha hills of the Western Ghats at a distance of 120 km upstream of the Ashtamudi estuary. The estuary has a maximum depth of over 6 m and is the deepest estuary in the State of Kerala. The annual average rainfall prevailing in the area is about 2400 mm and the estuary acts as flood storage for the city of Kollam. The estuary is rich in various mangrove species like *Avicennia officinalis* and supports over 57 species of avifauna. The estuary is a habitat for phytoplankton of different genera.

The estuary is stressed with anthropogenic activities including over-fishing, solid waste dumping (Priya et al. 2023), waste water discharge and cage fish farming. Being a hot tourist spot, the water quality is affected adversely due to unscientific practices in the catchment area. The excessive usage of chemical fertilizers in agricultural areas near the estuary has led to the inflow of nutrients, leading to eutrophication. The estuary has been classified as eutrophic as reported in previous studies (Sruthy et al. 2021). This necessitates the

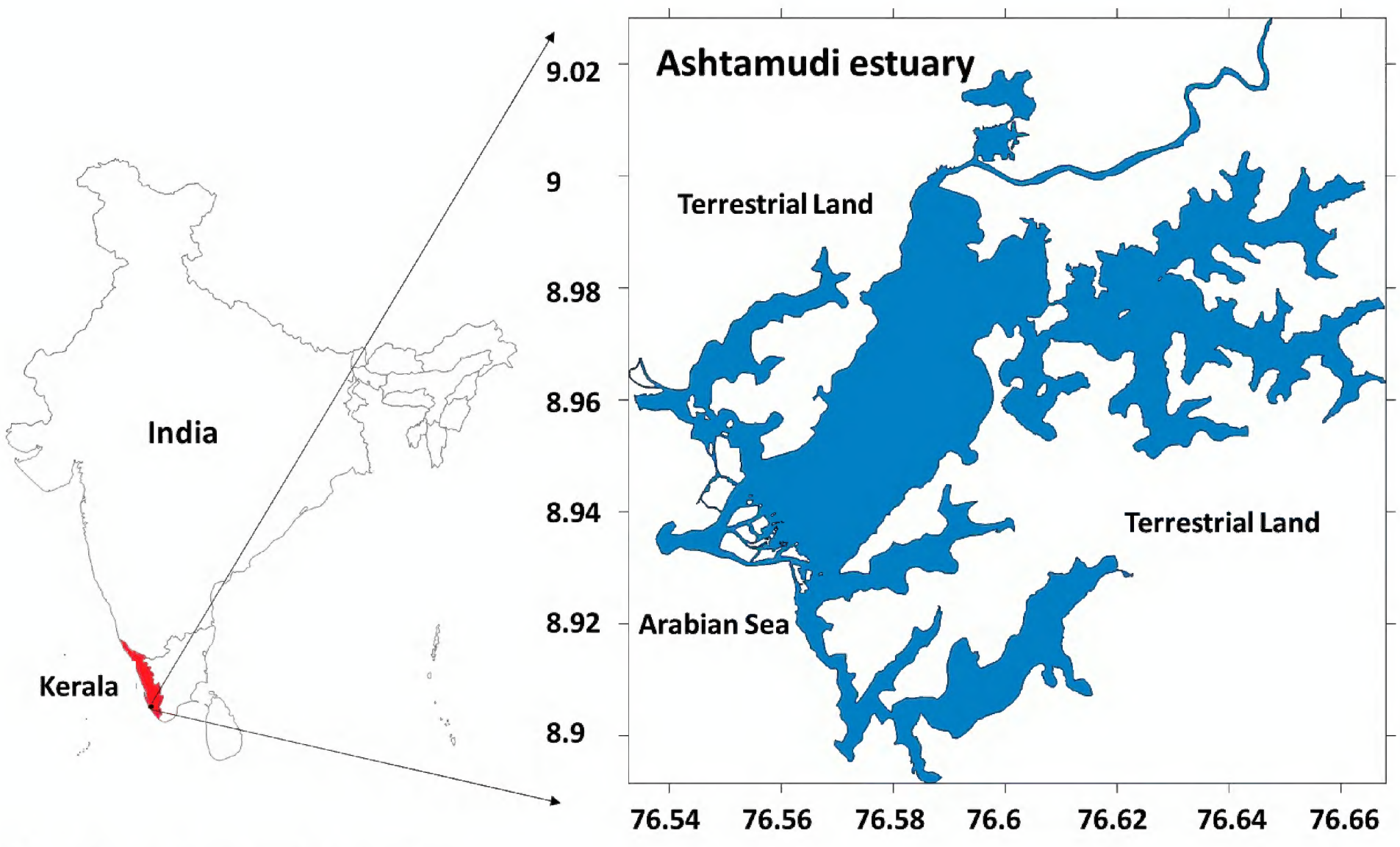


**Figure 1.** Location map of the study area.

development of prediction models for the most prominent parameter that represents eutrophication state namely, chlorophyll-a. For this, data pertaining to various water quality parameters such as pH, turbidity, salinity, DO, TN, TP, chlorophyll-a were collected from the government agencies during the periods 2012 to 2018 over a monthly basis. The location of sample station was at mid-estuarine reaches and the data from one single station alone could be considered for the present study due to lack of data from spatial stations. However, the data analysis of one station shall help in identifying the effect of salinity on chlorophyll-a at a particular location, and hence, salinity can be used as an indirect parameter for representing estuarine mixing. The raw data were cleaned and checked for outliers and then normalized to enhance the performance of the AI models. The data were set back to its original form after network training.

## 2.2 Identification of influential parameters

The performance of prediction models is dependent on the input variables. Thus, the significant variables affecting chlorophyll-a was selected appropriately for enabling the models to accurately predict chlorophyll-a concentration. In order to identify the significant influencing parameters, correlation coefficient between chlorophyll-a and water quality variables were estimated. The parameters that have exhibited a significant correlation of at least by 95% confidence level were selected as input parameters for model development.

## 2.3 Modelling chlorophyll-a

The chlorophyll-a prediction models were developed using Artificial Neural Networks (ANN) and Adaptive Neuro Fuzzy Inference System (ANFIS) modelling approaches. Of the total data set, 70% of the data were used for training the models and the remaining 30% of the data were used for testing and validation purposes.

### 2.3.1 ANN models

In the present study, three ANN models were developed and tested: Back Propagation Neural Network (BPNN), Generalized Regression Neural Network (GRNN), and Recurrent Neural Network (RNN). BPNN is among the most commonly adopted NNs for solving environmental problems. It generally consists of an input layer that contains the input parameters, a hidden layer that has several neurons, the number of which is fixed using a trial-and-error procedure, and an output layer that contains the output variable, chlorophyll-a. The error in the prediction of chlorophyll-a is minimized by adjusting the weights during the training process and is done by distributing the errors using the back propagation algorithm (Tian et al. 2017). The BPNN architecture used for the study is shown in Fig 2a. GRNN is a type of radial basis function (RBF) neural network that has fast learning and rapid convergence capabilities (Chen and Liu 2015). They are commonly used to deal with non-linear problems. They consist of a feed-forward network having an input layer, one hidden layer and one output layer. The common GRNN formulation can be expressed as

$$y = \sum_{i=1}^{n} w_i \varphi_i(x) \quad (1)$$

Where $y$ is the output, $w_i$ is the weight between hidden and output layer, $\varphi_i(x)$ is the output value of the neuron in the hidden layer after transfer by GRNN.

RNN is a type of deep learning neural network model that works through parameter sharing and state transformation. It consists of an input and output layer and node groups, collectively called RNN cell. The model parameters are updated during training phase. A typical structure of RNN adopted for the study is shown in Fig. 2b.

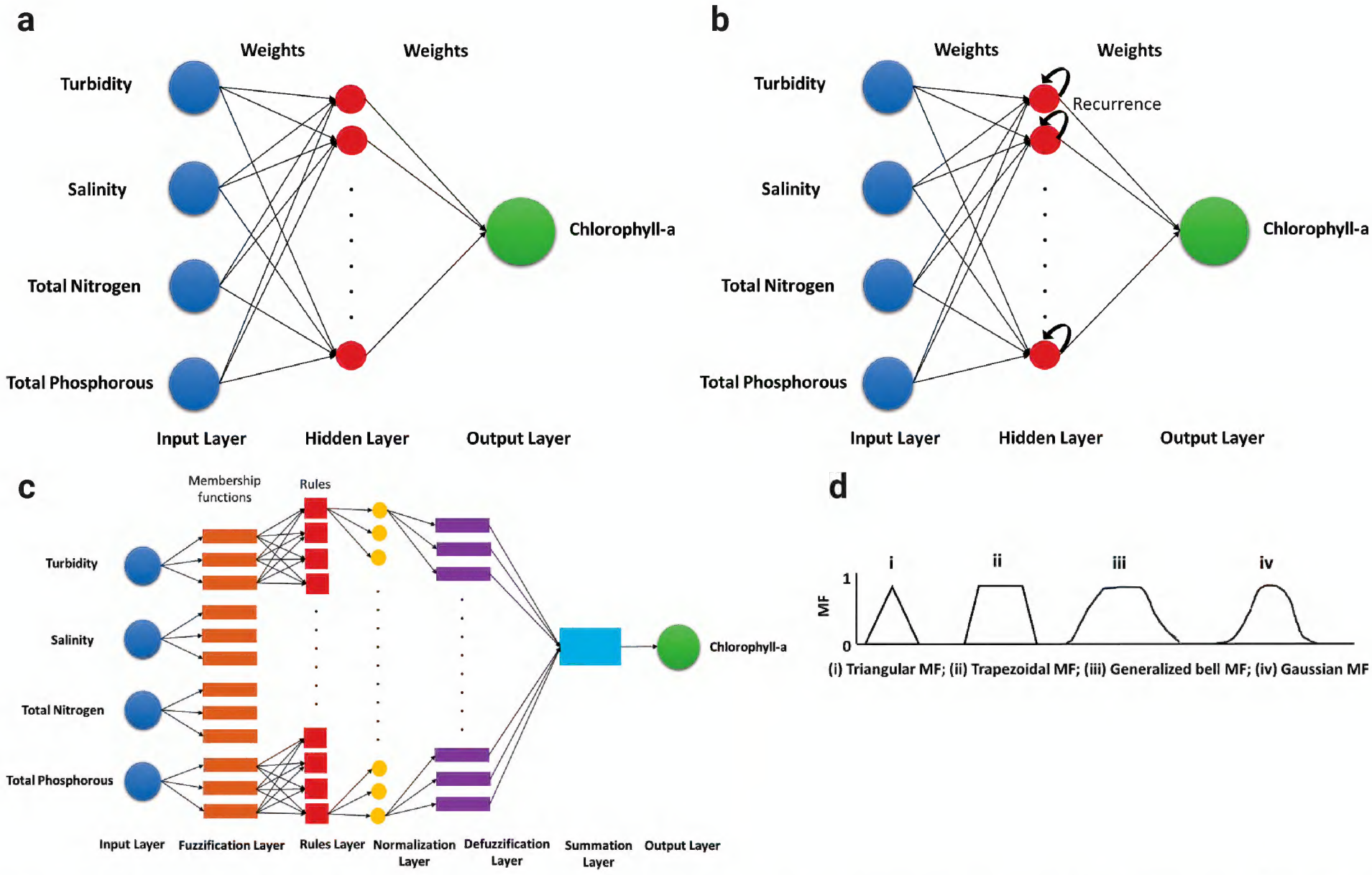


**Figure 2.** (**a**) BPNN architecture; (**b**) RNN architecture; (**c**) ANFIS architecture; and (**d**) membership functions for ANFIS model adopted for the study.

### 2.3.2 ANFIS models

A combination of ANN and fuzzy logic, ANFIS consists of a fuzzy sugeno system and multilayer feed forward system in a single framework. It is capable to solve complex non-linear problems in less computational period (Luo et al. 2019). A typical ANFIS consists of six layers: an input layer, a fuzzification layer, a rules layer, a normalized layer, a defuzzification layer and an output layer. The two parameters that are used in ANFIS include premise parameters that are modified using back propagation during training phase and found in the fuzzification layer, and consequent parameters that are adjusted by method of least squares during the training phase and found in the defuzzification layer. The general architecture of ANFIS adopted for the study is represented in Fig. 2c. The type of membership function selected and the rules developed are crucial steps involved in ANFIS modelling. The various types of membership functions adopted for the study include triangular (Tri MF), trapezoidal (Trap MF), Generalized bell (Gbell MF) and Gaussian (Gauss MF) as shown in Fig. 2d.

### 2.4 Performance evaluation of models

In order to identify the optimal chlorophyll-a prediction model, the performance indicators namely, coefficient of determination ($R^2$) and root-mean-square error (RMSE) were used. The higher $R^2$ and lower RMSE were used as the criteria for identifying the best prediction models for chlorophyll-a.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(x_i - y_i)^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \quad (2)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - x_i)^2} \quad (3)$$

Where n is the number of observations, $x_i$ is the observed chlorophyll-a concentration, $y_i$ is the predicted chlorophyll-a concentration, $\bar{x}$ is the mean value of observed chlorophyll-a.

## 3. Results and discussion

### 3.1 Water quality parameters

The water quality parameters namely, pH, turbidity, salinity, dissolved oxygen, total nitrogen, total phosphorous and chlorophyll-a were used for the study. Table 1 gives the statistical summary of the parameters. The turbidity variations were between 0.251 and 19.98 NTU, a higher turbidity can indirectly represent higher chlorophyll-a concentration. Salinity variations were much higher in the estuarine environment and is much dependent on the freshwater flow and tidal flow. A minimum of 1.25‰ was observed during monsoon period, while salinity as high as 30‰ has been detected during the pre-monsoon, when the freshwater influx was negligible. There was a wide variation in pH and dissolved oxygen, the estuarine waters were predominantly alkaline in nature with satisfactory to good dissolved oxygen for supporting aquatic life. The mixing and circulation within the estuarine reaches has been the main reason for saturated dissolved oxygen in the estuary, except for a depletion during pre-monsoon period. The total nitrogen and total phosphorous varied from 110 to 950 µg/L and 15 to 490 µg/L respectively, with higher values observed during pre-monsoon season. A mean value of 7.6 µg/L of chlorophyll-a was observed with variations between 0.9 and 17.9 µg/L over a temporal scale.

**Table 1.** Statistical summary of water quality parameters.

| Parameter | Minimum | Maximum | Mean | Std Dev |
|---|---|---|---|---|
| **Turbidity (NTU)** | 0.251 | 19.98 | 5.06 | 5.53 |
| **Salinity (‰)** | 1.25 | 30.0 | 21.735 | 12.253 |
| **pH** | 6.3 | 8.7 | 7.9 | 1.4 |
| **Dissolved Oxygen (mg/L)** | 4.1 | 8.2 | 6.5 | 2.5 |
| **Total Nitrogen (µg/L)** | 110 | 950 | 410 | 130 |
| **Total Phosphorous (µg/L)** | 15 | 490 | 240 | 64 |
| **Chlorophyll a (µg/L)** | 0.9 | 17.9 | 7.6 | 2.7 |

## 3.2 Identification of input parameters

The chlorophyll-a variations are influenced by many parameters like availability of sunlight, nutrients, mixing conditions and these parameters are represented by measurable variables like turbidity, salinity, total nitrogen and total phosphorous. The concentration of chlorophyll-a can also affects the water quality parameters such as pH and dissolved oxygen or in other terms, the variations in these parameters can indirectly represent the levels of chlorophyll-a. Therefore, to identify the factors influencing chlorophyll-a, its correlation with the water quality parameters were estimated using correlation coefficient, which is a commonly adopted indicator of the relation between two parameters. Table 2 represents that chlorophyll-a has the maximum correlation with total nitrogen and total phosphorous, with a higher significance of 99% confidence level, followed by turbidity and salinity, whose correlation with chlorophyll-a is significant at 95% confidence level. Dissolved oxygen was observed to have a negative weak correlation, while pH had a weak positive correlation, but were not significant. Accordingly, total nitrogen, total phosphorous, turbidity and salinity were considered for developing prediction models for chlorophyll-a.

**Table 2.** Correlation coefficient of chlorophyll a with water quality parameters.

| | Turbidity | Salinity | pH | Dissolved Oxygen | Total Nitrogen | Total Phosphorous |
|---|---|---|---|---|---|---|
| **Chlorophyll a** | 0.77* | 0.73* | 0.529 | -0.632 | 0.796** | 0.893** |

* represents significance at $p < 0.05$;
** represents significance at $p < 0.01$.

## 3.3 Prediction models for chlorophyll-a

The water quality parameters namely, turbidity, salinity, total nitrogen and total phosphorous showed significant correlation with chlorophyll-a. It is well recognized that total nitrogen and total phosphorous are causative parameters of chlorophyll-a, while turbidity can be considered as a response variable, because higher chlorophyll concentrations lead to the enhancement in water turbidity levels. Many studies that have developed prediction models for chlorophyll-a have used cause variables like nitrogen and phosphorous (Vollenweider et al. 1998; Gupta 2014; Luo et al. 2019) and water quality parameters that act as response variables too (Alves et al. 2013; O'Boyle et al. 2013; Chen and Liu 2015; Santos 2015). However, the salinity has been identified to have a significant correlation with chlorophyll-a in the Ashtamudi estuary. Therefore, to analyse the effect of salinity on the prediction of chlorophyll-a, two scenarios have been considered:

- Scenario I:
  Chl-a = f(Turb, TN, TP)
- Scenario II:
  Chl-a = f(Turb, Sal, TN, TP)

Where Chl-a is the chlorophyll-a concentration, Turb is the turbidity, Sal is the salinity, TN is the total nitrogen and TP is the total phosphorous. In order to further identify the best fit model, seven models were trained and tested using the avail-

able data: three neural network models namely, BPNN, GRNN, RNN; and four ANFIS models with triangular, trapezoidal, Gbell and Gaussian membership functions. Table 3 gives the model performance indicators for various models. Fig. 3 shows the predictability of chlorophyll-a using various models for the two scenarios considered for the study. It is observed that of the two scenarios, the one which considered salinity along with cause and response variables has improved the predictability of chlorophyll-a. GRNN model has been identified to be best suited for predicting chlorophyll-a for both the scenarios; by incorporating salinity, the predictability has enhanced marginally, leading to a prediction with an $R^2$ of 0.9976 for training and 0.9982 for testing, nearly approaching 1. The RNN model was identified to have the worst performance in predicting chlorophyll-a for scenario I with an $R^2$ of 0.37 and 0.21 respectively while training and testing and thus exhibiting drastic enhancement in its performance. Similarly, the performance of BPNN model was very poor for scenario I, especially during the testing phase, with an RMSE as high as 0.85, while its performance considerably improved for scenario II with 90.6% reduction in RMSE. These results highlight that ANN and RNN are much sensitive to the input parameters, while GRNN is more stable and can be applied in estuarine environments at various salinity conditions with less sensitiveness for the predicted chlorophyll-a with respect to the input variables. However, mostly all the ANFIS models were likely sensitive to the input variables almost to the same degree. Of the four ANFIS models, the models with Gaussian MF and Generalized bell MF exhibited good performance in chlorophyll-a predictions, especially for scenario II, followed by trapezoidal and triangular MF models. Surprisingly, they showed higher $R^2$ during testing phase than training phase.

**Table 3.** Performance evaluation of prediction models.

| Prediction Models | Scenario I | | | | Scenario II | | | |
|---|---|---|---|---|---|---|---|---|
| | Training | | Testing | | Training | | Testing | |
| | $R^2$ | RMSE | $R^2$ | RMSE | $R^2$ | RMSE | $R^2$ | RMSE |
| **BPNN** | 0.63 | 0.13 | 0.06 | 0.85 | 0.73 | 0.09 | 0.73 | 0.08 |
| **GRNN** | 0.81 | 0.02 | 0.77 | 0.01 | 0.99 | 0.01 | 0.99 | 0.01 |
| **RNN** | 0.37 | 0.64 | 0.21 | 0.59 | 0.99 | 0.01 | 0.73 | 0.09 |
| **ANFIS1** | 0.57 | 0.12 | 0.82 | 0.08 | 0.86 | 0.07 | 0.93 | 0.02 |
| **ANFIS2** | 0.57 | 0.13 | 0.85 | 0.07 | 0.84 | 0.08 | 0.97 | 0.01 |
| **ANFIS3** | 0.66 | 0.11 | 0.88 | 0.06 | 0.93 | 0.03 | 0.98 | 0.01 |
| **ANFIS4** | 0.66 | 0.1 | 0.89 | 0.06 | 0.94 | 0.03 | 0.98 | 0.01 |

Note: Scenario I: chl a = f(Turb, pH, DO); Scenario II: chl a = f(Turb, pH, DO, Sal); Scenario III: chl a = f(TN, TP); Scenario IV: chl a = f(Turb, TN, TP); Scenario V: chl a = f(Turb, Sal, TN, TP).
ANFIS1: Triangular MF; ANFIS2: Trapezoidal MF; ANFIS3: Gbell MF; ANFIS4: Gauss MF.

Fig. 4 shows the validation of various models for the two scenarios studied. It is evident that ANN model over-predicted chlorophyll-a values at low chlorophyll-a levels, while under-predicted at higher chlorophyll-a levels for scenario I, while this trend was altered for scenario II, wherein the over-prediction of chlorophyll-a was considerably reduced at lower levels. In the case of GRNN, the model marginally under-predicted chlorophyll-a concentration at higher concentration levels, while there was a perfect fit of predicted chlorophyll-a with the observed values for scenario II. Even though there were wide under-predictions in chlorophyll-a for the RNN model, a good fit was obtained for scenario II. The predictability of all of the ANFIS models was more or less similar for

scenario I with chlorophyll-a under-predictions at higher concentration, while the inclusion of salinity as an input parameter enabled the models to accurately predict the observed chlorophyll-a concentrations, thereby highlighting the role of salinity in the prediction of chlorophyll-a in estuarine environments.

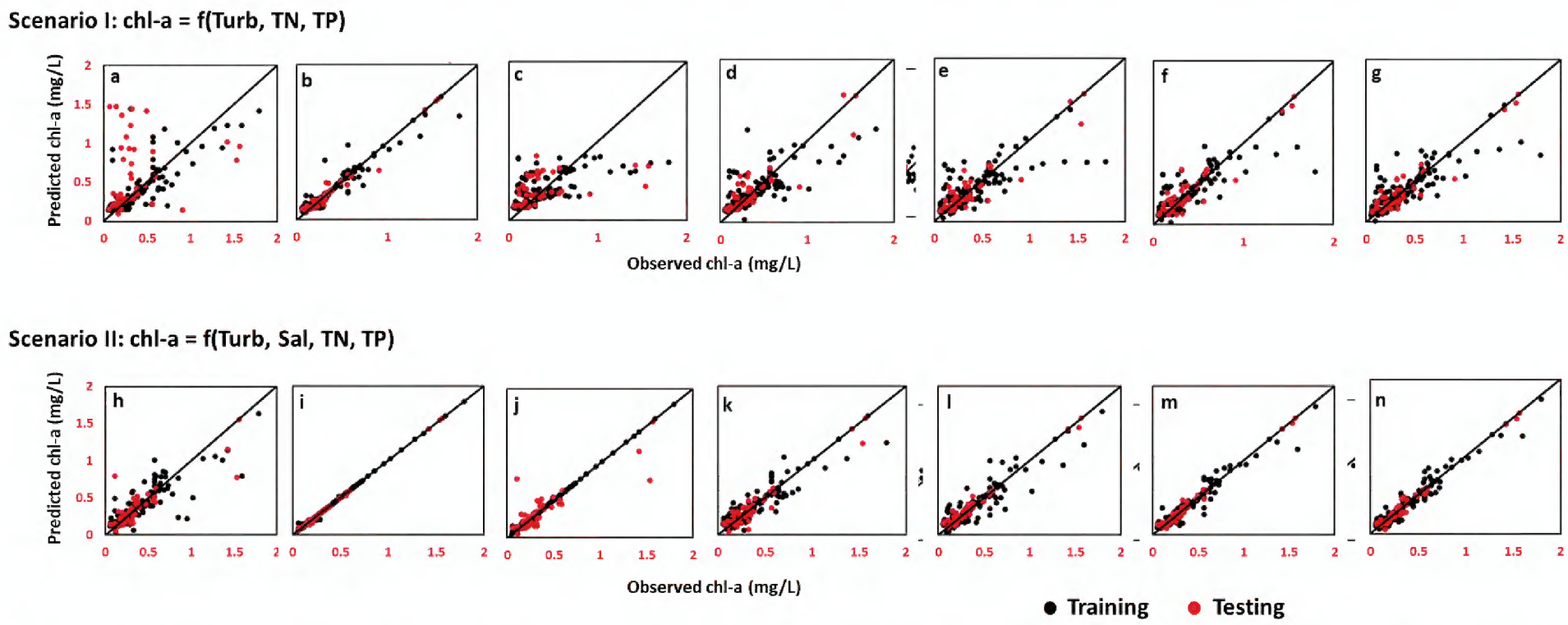


**Figure 3.** Predicted chlorophyll-a for scenario I using (**a**) BPNN; (**b**) GRNN; (**c**) RNN; (**d**) ANFIS – triangular MF; (**e**) ANFIS – trapezoidal MF; (**f**) ANFIS – Gbell MF; (**g**) ANFIS – Gauss MF; and for scenario II using (**h**) GPNN; (**i**) GRNN; (**j**) RNN; (**k**) ANFIS – triangular MF; (**l**) ANFIS – trapezoidal MF; (**m**) ANFIS – Gbell MF; (**n**) ANFIS – Gauss MF.

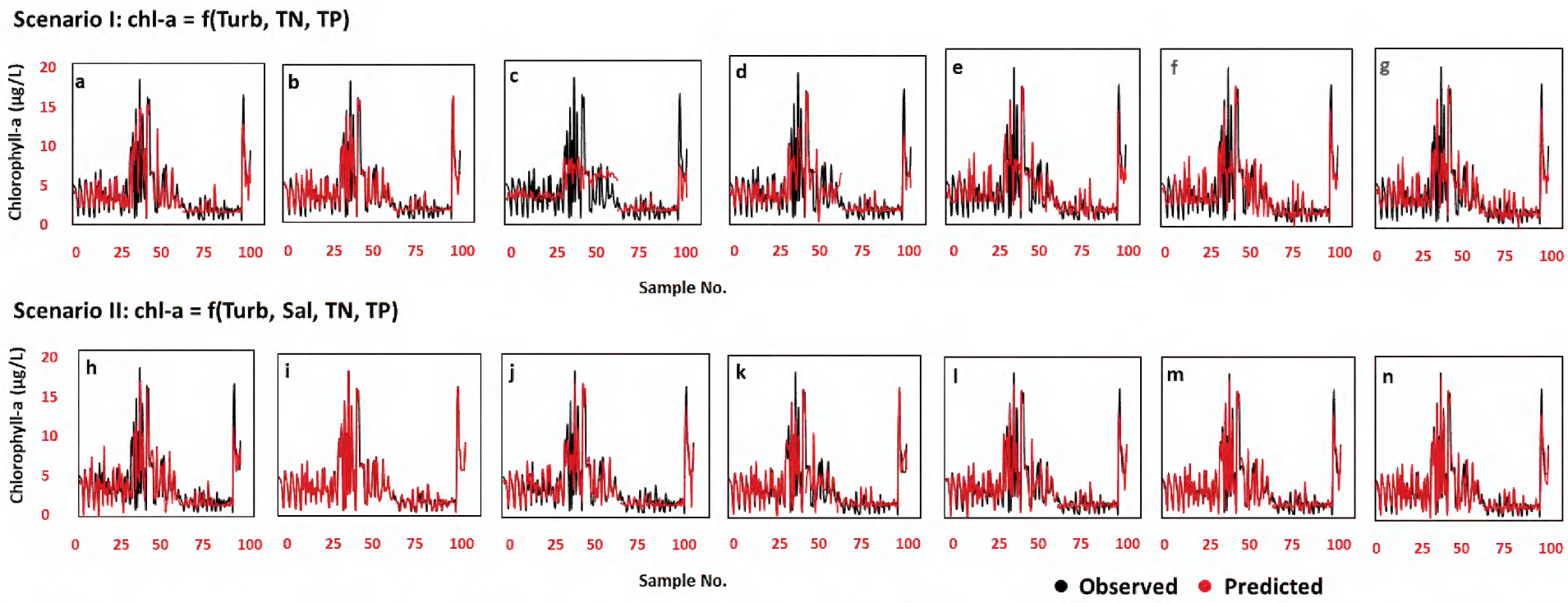


**Figure 4.** Validation of predicted chlorophyll-a for scenario I using (**a**) BPNN; (**b**) GRNN; (**c**) RNN; (**d**) ANFIS – triangular MF; (**e**) ANFIS – trapezoidal MF; (**f**) ANFIS – Gbell MF; (**g**) ANFIS – Gauss MF; and for scenario II using (**h**) GPNN; (**i**) GRNN; (**j**) RNN; (**k**) ANFIS – triangular MF; (**l**) ANFIS – trapezoidal MF; (**m**) ANFIS – Gbell MF; (**n**) ANFIS – Gauss MF.

### 3.4 Impact of estuarine mixing on chlorophyll-a

The salinity was identified to play a crucial role in chloroplyll-a dynamics in the Ashtamudi estuary. The incorporation of salinity as an input parameter significantly enhanced the predictability of chlorophyll-a. The variations of chlorophyll-a were observed to have a diversity at various salinity values as shown in Fig. 5a. The variations of chlorophyll-a were rather more pronounced at higher salinity values. At salinity values less than 5‰, there was an overall decreasing trend for chlorophyll-a with salinity, but the relation was not much significant (Fig. 5b). However, when the salinity ranged between 5 and 30‰, there was a pronounced increase

in chlorophyll-a with an increase in salinity. As a larger % of the data set fall in the higher salinity category (of >5‰), the correlation between salinity and chlorophyll-a was obtained as positive. Even though the chlorophyll-a had a non-linear variation with salinity, the AI models were capable to replicate these variations and was able to predict the chlorophyll-a concentration at all salinity levels. At lower salinity levels, the freshwater flow is appreciable and this will create salinity gradients, causing the estuarine reaches to act as partially-stratified or stratified zones. In these zones, mixing is relatively less. A higher near-surface salinity represents higher mixing in the water column and these zones act as well-mixed zones. Higher mixing can promote the penetration of light into deeper zones, thereby enhancing primary productivity. This will have implications on the enhancement of chlorophyll-a concentrations. Thus, the study brings out the hypothesis that well-mixed conditions in estuarine reaches are associated with higher chlorophyll-a concentration and that salinity can be used as an indirect representation of mixing conditions that plays crucial role for chlorophyll-a prediction. Studies have to be extended to analyse the performance of the models on a temporal and spatial basis and for various tidal conditions. Further, an in-depth analysis on the influence of other water quality parameters on the chlorophyll-a dynamics and prediction is also to be performed. The validation of this hypothesis has to be extended to spatial attributes of the Ashtamudi estuary and also needs validation in other world estuaries.

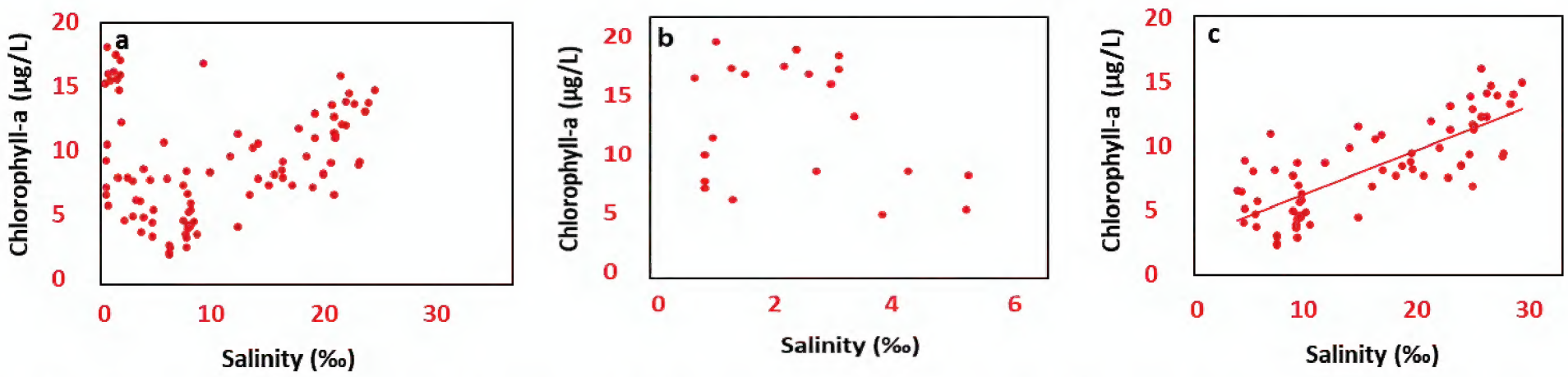


**Figure 5.** Variation of chlorophyll-a with salinity (**a**) for salinity between 0 and 30‰; (**b**) for salinity between 0 and 5‰; and (**c**) for salinity between 5‰ and 30‰.

## 4. Conclusions

The study evaluates the influence of freshwater seawater mixing on the chlorophyll-a prediction in the Ashtamudi estuary on the south-west coast of India. The water quality parameters such as total nitrogen, total phosphorous, turbidity and salinity exhibit significant correlation with chlorophyll-a, thereby indicating that in estuarine mixing environment, salinity play a significant role in chlorophyll-a fluctuations. The chlorophyll-a variations are more prominent at salinity values greater than 5‰ and is an increasing function of salinity. The chlorophyll-a is thus influenced by the mixing regimes pertaining in the estuary and salinity can possibly indirectly represent the mixing conditions of the estuary. Of the various ANN models, GRNN can very well be used for predicting chlorophyll-a concentrations in estuaries. Among the ANFIS models, the selection of membership function play crucial role in the predictability. Gaussian and Generalized bell membership function-based ANFIS models outperformed those with triangular and trapezoidal membership functions. Overall, ANFIS can very well be used for predicting chlorophyll-a due to their non-linear be-

haviour. The effect of salinity on chlorophyll-a prediction models need testing in other estuarine environments and for other modelling approaches to validate and generalize the findings of the present study.

## Additional information

### Conflict of interest

The authors have declared that no competing interests exist.

### Ethical statement

No ethical statement was reported.

### Funding

No funding was reported.

### Author contributions

Conceptualization: KLP. Data curation: SH. Formal analysis: JCCC. Investigation: SH, KLP. Writing - original draft: JCCC. Writing - review and editing: KLP, SH.

### Author ORCIDs

K. L. Priya https://orcid.org/0000-0002-8433-2920
Joan Cecilia C. Casila https://orcid.org/0000-0001-6319-8999

### Data availability

All of the data that support the findings of this study are available in the main text.

## References

Alves G, Flores-Montes M, Gaspar F, Gomes J, Feitosa F (2013) Eutrophication and water quality in a tropical Brazilian estuary. Journal of Coastal Research 65: 7–12. https://doi.org/10.2112/SI65-002.1

Amorim F, Rick, J, Lohmann G, Wiltshire KH (2021) Evaluation of machine learning predictions of a highly resolved time series of chlorophyll-a concentration. Applied Sciences 11(16): 7208. https://doi.org/10.3390/app11167208

Cen H, Jiang J, Han G, Lin X, Liu Y, Jia X, Ji Q, Li B (2022) Applying Deep Learning in the prediction of chlorophyll-a in the East China Sea. Remote Sensing 14(21): 5461. https://doi.org/10.3390/rs14215461

Chang CF, Garcia, V, Tang C, Vlahos P, Wanik D, Yan J, Bash JO, Astitha M (2021) Linking multi-media modeling with machine learning to assess and predict lake chlorophyll a concentrations. Journal of Great Lakes Research 47(6): 1656–1670. https://doi.org/10.1016/j.jglr.2021.09.011

Chen W-B, Liu W-C (2015) Water quality modeling in reservoirs using multivariate linear regression and two neural network models. Advances in Artificial Neural Systems 2015(1): 521721. https://doi.org/10.1155/2015/521721

Gayathri H, Sreelekshmi S, Gowtham Mohan, Gopika Sankar MS, Priya KL, Azhikodan G, Yokoyama K (2022) Association of eutrophication parameters and water quality parameters of two selected estuaries. Proceedings of the 19th Annual Meeting of the Asia Oceania Geosciences Society (AOGS 2022), August 1–5, 2022. https://doi.org/10.1142/9789811275449_0045

Gupta M (2014) A new trophic state index for lagoons. Journal of Ecosystems 2014(1): 152473. https://doi.org/10.1155/2014/152473

Huang Y, Pan J, Devlin AT (2023) Enhanced estimate of chromophoric dissolved organic matter using machine learning algorithms from Landsat-8 OLI data in the Pearl River Estuary. Remote Sensing 15(8): 1963. https://doi.org/10.3390/rs15081963

Kim HR, Soh HY, Kwak MT, Han SH (2022) Machine learning and multiple imputation approach to predict chlorophyll-a concentration in the coastal zone of Korea. Water 14(12): 1862. https://doi.org/10.3390/w14121862

Kim KM, Ahn JH (2022) Machine learning predictions of chlorophyll-a in the Han River basin, Korea. Journal of Environmental Management 318: 115636. https://doi.org/10.1016/j.jenvman.2022.115636

Luo W, Zhu S, Wu S, Dai J (2019) Comparing artificial intelligence techniques for chlorophyll-a prediction in US lakes. Environmental Science and Pollution Research 26: 30524–30532. https://doi.org/10.1007/s11356-019-06360-y

Mozo A, Morón-López J, Vakaruk S, Pompa-Pernía ÁG, González-Prieto Á, Aguilar JAP, Gómez-Canaval S, Ortiz JM (2022) Chlorophyll soft-sensor based on machine learning models for algal bloom predictions. Scientific Reports 12: 13529. https://doi.org/10.1038/s41598-022-17299-5

O'Boyle S, McDermott G, Noklegaard T, Wilkes R (2013) A simple index of trophic status in estuaries and coastal bays based on measurements of pH and dissolved oxygen. Estuaries and Coasts 36: 158–173. https://doi.org/10.1007/s12237-012-9553-4

Priya KL, Jegathambal P, James EJ (2015) On the factors affecting the settling velocity of fine suspended sediments in a shallow estuary. Journal of Oceanography 71(2): 163–175. https://doi.org/10.1007/s10872-014-0269-x

Priya KL, Shabana I, Gopika B, Michi M, Haddout S, Athul MM (2023) Implications of solid waste dumps on the microplastic abundance in groundwater in Kollam, India. Journal of Environmental Management 348: 119224. https://doi.org/10.1016/j.jenvman.2023.119224

Santos SJC (2015) [Personal Communication] Determination of the Trophic State to from the Application of Indices O'Boyle and TRIX in the Compartments of the Bay of Guanabara. RJ. Dissertation (Master in Geosciences (Geochemistry)) - Fluminense Federal University.

Savoy P, Harvey JW (2023) Predicting daily river chlorophyll concentrations at a continental scale. Water Resources Research 59(11): e2022WR034215. https://doi.org/10.1029/2022WR034215

Song J, Jiang W, Xin L, Zhang X (2024) Predicting the temporal-spatial distribution of chlorophyll-a in the Yellow River Estuary using explainable machine learning. Estuarine, Coastal and Shelf Science 304: 108820. https://doi.org/10.1016/j.ecss.2024.108820

Sruthy GS, Priya KL, Athul MM, Suchith C, Adarsh S, Haddout S (2021) Fuzzy logic approach for the assessment of trophic state of water bodies. Ecological Engineering 169: 106314. https://doi.org/10.1016/j.ecoleng.2021.106314

Tian W, Liao Z, Zhang J (2017) An optimization of artificial neural network model for predicting chlorophyll dynamics. Ecological Modelling 364: 42–52. https://doi.org/10.1016/j.ecolmodel.2017.09.013

Vollenweider RA, Giovanardi F, Montanari G, Rinaldi A (1998) Characterization of the trophic conditions of marine coastal waters with special reference to the NW Adriatic Sea: proposal for a trophic scale, turbidity and generalized water quality index. Environmetrics 9: 329–357. https://doi.org/10.1002/(SICI)1099-095X(199805/06)9:3%3C329::AID-ENV308%3E3.0.CO;2-9